

Predicting Hearing Loss From Otoacoustic Emissions Using An Artificial Neural Network

Rouviere de Waal, René Hugo, Maggi Soer and Johann J Krüger

Department of Communication Pathology, & Electronic Engineering
University of Pretoria,

ABSTRACT

Normal and impaired pure tone thresholds (PTTs) were predicted from distortion product otoacoustic emissions (DPOAEs) using a feed-forward artificial neural network (ANN) with a back-propagation training algorithm. The ANN used a map of present and absent DPOAEs from eight DPgrams, ($2f_1 - f_2 = 406 - 4031$ Hz) to predict PTTs at 0.5, 1, 2 and 4 kHz. With normal hearing as < 25 dB HL, prediction accuracy of normal hearing was 94% at 500, 88% at 1000, 88% at 2000 and 93% at 4000 Hz. Prediction of hearing-impaired categories was less accurate, due to insufficient data for the ANN to train on. This research indicates the possibility of accurately predicting hearing ability within 10 dB in normal hearing individuals and in hearing-impaired listeners with DPOAEs and ANNs from 500 - 4000 Hz.

Key Words: Otoacoustic emissions, Distortion product otoacoustic emissions, Artificial neural networks, Hearing threshold prediction, Objective hearing assessment.

INTRODUCTION

David Kemp (1978) first described otoacoustic emissions (OAE) from the human ear and ignited tremendous interest in measurements of these emissions to develop another objective diagnostic test of hearing. Distortion product otoacoustic emissions (DPOAEs) are relatively easy measurable sinusoids, recordable in the occluded ear canal during the simultaneous stimulation of two primary pure tone frequencies, f_1 and f_2 with $f_2 > f_1$. The current view on DPOAE generation is that these active responses from the cochlea have two main sources of energy (Knight & Kemp, 1999; Mauermann, Uppenkamp, Van Hengel, & Kollmeier, 1999a; Mauermann, Uppenkamp, Van Hengel, & Kollmeier, 1999b; Talmadge, Long, Tubis, & Dhar, 1998). The first primary source of DPOAE energy is the result of nonlinear interaction between the two primary frequencies on the basilar membrane at the f_2 place, also referred to as the generation site. The second source of DPOAE energy is caused by the reflection of the coherent waves at the $2f_1 - f_2$ frequency place, also referred to as the re-emission site. The measured DPOAE in the ear canal is the result of the interference of both these sources. It has also recently been postulated that there are two mechanisms responsible for DPOAE generation:

DPOAEs consist of a mixture of nonlinear energy arising from two locations on the basilar membrane as well as linear coherent reflection off pre-existing micromechanical impedance perturbations (Kalluri & Shera, 2001).

The correlation between DPOAEs and hearing sensitivity has kept many researchers occupied in the last two decades (Bonfils, Avan, Londero, Trotoux, & Narcy, 1991; Harris & Probst, 1991; Kimberley & Nelson, 1989; Kummer, Janssen, & Arnold, 1998; Martin, Ohlms, Franklin, Harris, & Lonsbury-Martin, 1990; Nieschalk,

Hustert, & Stoll, 1998; Probst & Hauser, 1990; Smurzynski, Leonard, Kim, Lafreniere, Marjorie, & Jung, 1990). The quest to predict pure tone thresholds (PTTs) accurately with DPOAEs arises not from the need to replace existing conventional behavioral evaluation procedures, but to aid in the assessment of pure tone sensitivity in difficult-to-test populations such as neonates, infants, malingerers and the critically ill. To determine PTTs in special populations with objective physiologic measurements such as tympanometry, the acoustic reflex, ABR and OAEs, procedures are often costly, require a large amount of time (Northern, 1991), highly trained and specialized personnel and sometimes involve sedation (Musiek, Berenstein, Hall III, & Schwaber, 1994). Above all, current objective procedures such as ABR, have a limited frequency area in which hearing sensitivity can be determined accurately (Weber, 1994). There is therefore a definite need for an objective, reliable, rapid and economic test of hearing that evaluates hearing sensitivity across a wide range of frequencies to aid in the assessment of difficult-to-test populations.

The main aims of most previous studies were attempts to categorize PTTs with DPOAEs as normal or impaired (Hurley & Musiek, 1994; Kimberley, Kimberley, & Roth, 1994a; Kimberley, Hernadi, Lee, & Brown, 1994b) or to gain more information regarding the site-of-lesion in diagnostic audiology (Moulin, Bera, & Collet, 1994; Ohlms, Lonsbury-Martin, & Martin, 1990; Robinette, 1992; Tanaka, O-Uchi, Arai, & Suzuki, 1987). Most researchers, however, found it extremely difficult or even impossible to predict impaired PTTs or to categorize hearing sensitivity at low frequencies as normal or impaired with DPOAEs (Gorga, Neely, Bergman, & Beauchaine, 1993; Kimberley et al., 1994b; Stover, Gorga, & Neely, 1996; Zhao & Stephens, 1998). This unsatisfactory prediction of PTTs with DPOAEs is probably due to the large number of DPOAE stimulus parameters that influence optimal measurement (Bonfils

et al., 1991; Gorga, et al., 1993), the complex non-linear nature of the measured responses (Kummer et al., 1998; Nakajima, Mountain, & Hubbard, 1998) and the inability of conventional statistics to address this problem sufficiently (Kimberley et al., 1994a).

Some of the previous studies that attempted to classify hearing sensitivity with DPOAEs as normal or impaired will be reviewed shortly. Kimberley and Nelson (1989) determined the correlation between DPOAE threshold and PTTs in 21 ears (11 normal, 10 with a degree of sensorineural hearing loss) using an f_2/f_1 ratio of 1.2 and measuring DPOAE input-output (I/O) functions from 30 - 80dB SPL in 6dB steps. They claimed that DPOAE thresholds predicted PTTs within 10dB over a range of sensory thresholds from 0 - 60dB SPL for the frequencies 700- 6000 Hz. This was the first report of such an accurate prediction. Kimberley et al. (1994b) predicted hearing status in normal and hearing-impaired ears with DPOAEs at six frequencies ranging from 1025 - 5712 Hz. The significance of variables such as DPOAE levels, age and gender were determined in the definition of normal versus abnormal PTTs and then applied to a new set of unfamiliar data to determine their predictive accuracy at each frequency. Classification accuracy of normal hearing varied from 71% at 1025 to 92% at 2050 Hz. Kimberley et al. (1994a) used an artificial neural network (ANN) approach to predict PTTs with DPOAEs and prediction accuracy varied from 57% correct classification of hearing impairment at 1025Hz to 100% at 2050Hz when normal hearing was defined as PTTs < 20dB HL. Overall classification accuracy was 80% for normal PTTs and 90% for impaired PTTs. Gorga et al. (1993) also measured the extent to which DPOAEs could accurately distinguish between normal-hearing and hearing-impaired ears. DPOAE levels at 65/55 dB SPL distinguished between normal and impaired subjects at 4000, 8000 and to a lesser extent at 2000 Hz. At 500 Hz, performance was no better than chance due to high biological noise levels such as breathing and swallowing. Moulin et al. (1994) categorized hearing as normal or impaired (below or above 25 dB HL) by comparing the DPOAE threshold with the PTT of the frequency that corresponded with the geometric mean of the primaries. Correct classification was 79.4% at 4000, 81.5% at 2826, 75.6% at 1413, 73.2% at 1000 and 52.9% at 706 Hz. False negative responses ranged from 11.9% to 16.6%.

From various reports it became clear that there are numerous DPOAE and demographic features that influence predictive accuracy of PTTs (Avan & Bonfils, 1993; Gaskill & Brown, 1990; Kimberley et al., 1994b, Mills, 1997; Moulin et al., 1994; Stover et al., 1996). Features mentioned in these studies include the DPOAE amplitude versus threshold correlation with PTTs, PTT frequency correlation to the frequency of f_1 , f_2 , $2f_1 - f_2$ or the geometric mean of the primaries, the level of stimuli to evoke DPOAEs and possible incorporation of DPOAE amplitudes of adjacent frequencies.

Moulin et al. (1994) reported that DPOAE threshold, rather than DPOAE amplitude, seemed to be the best parameter in predicting PTTs. Stover et al. (1996) argued that while DPOAE threshold offered a slightly better prediction of PTTs than DPOAE amplitude, it lengthened test times due to difficulty of threshold determination against a noisy background and therefore reduced clinical utility as a tool for identification of

hearing loss. Stover et al. (1996) and Kimberley et al. (1994b) observed that the single most important variable to categorize PTTs as normal or impaired was the DPOAE amplitude in response to moderate level primaries (L_1 at 55 or 60 dB SPL) with f_2 frequencies close to the PTT frequency. Mauermann et al. (1999b) found that the DPOAE fine structure (fine structure has been defined by Talmadge et al., 1998, as quasiperiodic variations in DPOAE amplitude and phase with variations in DPOAE frequency) might be a more sensitive indicator of hearing impairment than DPOAE amplitude alone. In cases where the primary frequencies fell in areas of normal hearing but the DPOAE frequency's corresponding PTT was impaired, the $2f_1 - f_2$ DPOAE was still measurable but the fine structure disappeared.

When it comes to the analysis of DPOAE data, there are different frequency variables of the DPOAE to use for the prediction of PTTs. Possibilities include the $2f_1 - f_2$ frequency, f_1 , f_2 or the geometric mean of the primaries. Some researchers correlated the geometric mean frequency of the primaries with the PTT frequency (Bonfils et al., 1991; Lonsbury-Martin & Martin, 1990; Martin et al., 1990). Others found that DPOAE amplitudes of frequencies at and adjacent to the f_2 frequency were most predictive of normal hearing sensitivity (Harris, Lonsbury-Martin, Stagner, Coats, & Martin, 1989; Kimberley et al., 1994a; Kimberley et al., 1994b; Kummer et al., 1998). Recent studies proved that the region close to the f_2 place on the basilar membrane is the primary source of DPOAE energy and therefore the f_2 frequency is currently the preferred frequency to correlate with pure tone thresholds. (Mauermann et al., 1999a; Mauermann et al., 1999b; Talmadge et al., 1998).

Regarding the level of the stimuli to use in measuring DPOAEs for PTT prediction purposes, most researchers agree that moderate and lower level stimuli (below 65 dB SPL) are more suitable for PTT prediction and that lower levels are more frequency specific (Avan & Bonfils, 1993; Gaskill & Brown, 1990; Kimberley et al., 1994b, Mills, 1997). According to Bonfils et al. (1991), when primary intensities higher than 60 dB SPL are used for stimulus generation, it is probable that only passive properties of the cochlea contribute to the emission and that it might not indicate true outer hair cell functioning. It has recently been postulated by Kalluri and Shera (2001) that as there are two mechanisms in the generation of DPOAEs, passive linear components may be present in even a low level DPOAE.

Age was one demographic variable that had a great influence on prediction accuracy of PTTs with DPOAEs (Kimberley et al., 1994a; Kimberley et al., 1994b; Lonsbury-Martin et al., 1990). Although many studies indicated that the negative correlation between DPOAE levels and age was due to changes in hearing sensitivity associated with aging, rather than age itself (Avan & Bonfils, 1993; He & Schmiedt, 1996; Karzon, Garcia, Peterein, & Gates, 1994; Nieschalk et al., 1998), researchers attempting to predict PTTs with DPOAEs suggested the inclusion of the age variable in the prediction scheme for more accurate predictions (De Waal, 1998; De Waal, 2000; Kimberley et al., 1994a; Kimberley et al., 1994b; Lonsbury- Martin, Cutler, & Martin, 1991).

Most studies that attempted to predict PTTs with DPOAEs used statistical methods such as multivariate discriminant analysis (Kimberley et al., 1994b), relative operating characteristic curves (Gorga et al., 1993), or routine statistical tests such as χ^2 , Student's *t* test, matched *t* tests and Pearson's correlation coefficients (Moulin et al., 1994). A study by Dorn, Piskorski, Gorga, Neely, & Keefe (1999) compared single-variable and multivariate statistical approaches in PTT prediction with DPOAEs and concluded that superior performance and greater predictive accuracy were obtained with multivariate techniques.

Kimberley et al. (1994a) used a different data processing technique called artificial neural networks (ANNs) for the prediction of PTTs with DPOAEs. This experiment compared the classification performance of ANNs with discriminant analysis and found that ANNs outperformed the more traditional statistical technique in the hearing-impaired cases. This was probably because of this algorithm's superior ability to model complex problems, determine nonlinear correlations and excellent predicting abilities (Blum, 1992; Rao & Rao, 1995). Blum (1992) identified three advantages of neural networks over conventional statistical methods that can be applied in the prediction of PTTs with DPOAEs: a) ANNs require less need to determine relevant factors a priori. Irrelevant data has such a low connection strength it has no effect on the outcome and the ANN determines which factors are relevant. b) The sophistication of the ANN model allows it to take hundreds of factors into account simultaneously and this directness of the model enables it to solve complex problems in much less time. c) ANNs are extremely fault tolerant and can learn on noisy or incomplete data, which is often the case with absent or noisy DPOAEs.

Kimberley et al. (1994a) categorized hearing as normal or impaired (normal was defined as less than 20dB HL) with DPOAEs and ANNs but made no attempt to categorize the magnitude of the hearing loss by performing any further predictions.

The aim for this study was to investigate if ANNs could categorize PTTs into different groups of hearing sensitivity, not by only distinguishing between normal and impaired, but to categorize impaired PTTs

into 10dB groups. ANNs were used to determine a correlation between selected measured variables of DPOAEs and PTTs and to apply the correlation to make a prediction. The measured variables included DPOAEs at eleven 2f1 - f2 frequencies ranging from 2f1 - f2 = 406 to 4031 Hz and PTT information at 0.5, 1, 2 and 4kHz. Controlled variables included the frequencies of the primaries, ranging from f1 = 500 - 5031 Hz with a fixed f2/f1 ratio of 1.2 and the levels of L1 and L2 ranging from L2 = 60 - 25 dB, L1 > L2 by 10 dB. There was experimentation with certain manipulated variables to determine their effect on PTT prediction accuracy. Manipulated variables included experimentation with DPOAE amplitude representation into the ANN: either as a categorical value (with the dummy variable technique, explained in data preparation) or as a fraction of the maximum amplitude value. Subject age was included in the ANN as a categorical value with the dummy variable technique and there was experimentation with age representation in 5- or 10-year categories. The subject variable, gender, was always included in the ANN and was depicted with a one or a zero. Other manipulated variables included ANN topology experimentation with middle level neuron quantities (80, 100 and 120) and error tolerance levels (0.1%, 0.2% and 0.3%). Middle level neuron quantities were doubled for experiments where amplitude was presented as a categorical value (referred to as ALT AMP) to balance the increase in inset neurons, as suggested by Rao and Rao (1995).

There is one major difference between this study and most other studies predicting PTTs with DPOAEs that should be clearly stated: This experiment did not use the single f2 value corresponding to the PTT frequency, but used the present and absent responses for all 11 f2 frequencies as ANN input in the prediction of each PTT frequency. The use of frequencies other than the single closest corresponding f2 frequency is not a new idea: Kimberley et al. (1994a) used amplitudes of DPOAE frequencies adjacent to the f2 frequency to correlate to the PTT frequency and found that it improved PTT prediction. Experimentation with the use of the whole spectrum of DPOAE information to predict a single PTT was attempted to enhance prediction abilities of low frequencies, especially 500 Hz, which has proven

Table 1: Distribution pattern for different types of hearing loss in the 120 ear data set.

	Ears Group 1: PTAs 0-15 dB HL	Ears Group 2: PTAs 16-35 dB HL	Ears Group 3: PTAs 36- 65dB HL
Flat audiogram Not more than 20dB variation between 0.5 - 4 kHz.	40	11	16
Gradual slope PTTs increased gradually as frequency increased	0	9	24
Ski-slope Flat configuration up to 2 kHz with >20dB PTT drop in high frequencies	0	10	1
Low frequency loss 0.5 - 1 kHz more impaired than 2 - 4 kHz	2	0	0
Notch Notch shaped loss around 1 - 3 kHz	4	3	0

difficult or impossible for many other researchers (Durrant, 1992; Gorga et al., 1993; Moulin et al., 1994; Probst & Hausser, 1990; Stover et al., 1996).

METHOD

Subject

Data was obtained from 70 subjects (120 ears, in some cases only one ear fell within subject selection specifications), 28 males and 42 females ranging from 8 to 82 years old. Subjects were divided into three groups. The first group represented ears with normal hearing and had pure tone averages (PTAs) of 0 to 15 dB HL. For the determination of PTAs, 500, 1000, 2000 and 4000 Hz were taken in consideration. Four thousand Hertz was also included since it was one of the frequencies to be predicted. The second group represented mild hearing loss and had PTAs between 16 and 35 dB HL and the third group represented moderate hearing loss with PTAs between 36 and 65 dB HL. There were 46 ears in the first group, 33 in the second and 41 in the third group. The patterns of hearing loss of this population are presented in Table 1. Age and gender distribution of subjects can be seen in Figure 1.

Sampling

Subjects were drawn from a private practice in Audiology, the clinic at the University of Pretoria, and a school for the hard of hearing. The main aims of the study and the procedure for obtaining data were described and subjects were asked if they would be willing to participate.

Subject selection criteria

The selection criterion was normal middle ear

functioning. Subjects demonstrating type A tympanograms with static immittance between 0.3 to 1.6 mmhos (measured at 226 Hz) and a peak (or point of maximum admittance) between -100 daPa to 25 daPa were accepted for this study.

There was no selection criterion regarding age or gender. Only the pediatric population was excluded from this study due to differences in middle ear properties such as canal length, canal volume and middle ear reverse transmission efficiency that may have caused differences in DPOAE amplitudes (Lee, Kimberley, & Brown, 1993).

Subject selection procedure

If a subject agreed to participate in the study, a brief interview obtained limited background information such as subject name, gender and date of birth, information regarding hearing status and family history of hearing loss, history of middle ear problems, noise exposure, tinnitus or vertigo and medications currently used. The interview was followed by otoscopic examination, tympanometry and pure tone audiometry. Pure tone audiometry was performed in a sound proof booth on a GSI 60 audiometer. When PTTs exceeded 10 dB HL, pure tone bone conduction was also performed. Threshold determination was in 5 dB steps and a threshold was defined as the lowest hearing level with a minimum of two out of three responses at the specified dB level (Yantis, 1994). If a subject met the subject selection criteria, DPOAE measurements followed in a quiet room. The GSI 60 audiometer and GSI 28-A tympanometer met calibration requirements and the DPOAE probes were calibrated for the quiet room in which testing was performed.

Specification of dpoae stimulus parameters

The frequency range evaluated spanned from $f_1 = 500$ to 5031 Hz with a f_2/f_1 ratio of 1.2. Three data points per octave were selected as per the Grason-Stadler Inc. DPOAE user manual (1997) for Model GSI-60, resulting in 11 primary frequency pairs.

In the present study, eight DPgrams were obtained for each ear. All 11 frequency pairs were presented to the subjects at different levels, starting with maximum levels at $L_1 = 70$ dB; $L_2 = 60$ dB. Levels were decreased in 5 dB steps. $L_1 > L_2$ by 10 dB for every DPgram. The lowest level for the primaries tested was $L_1 = 35$ dB; $L_2 = 25$ dB.

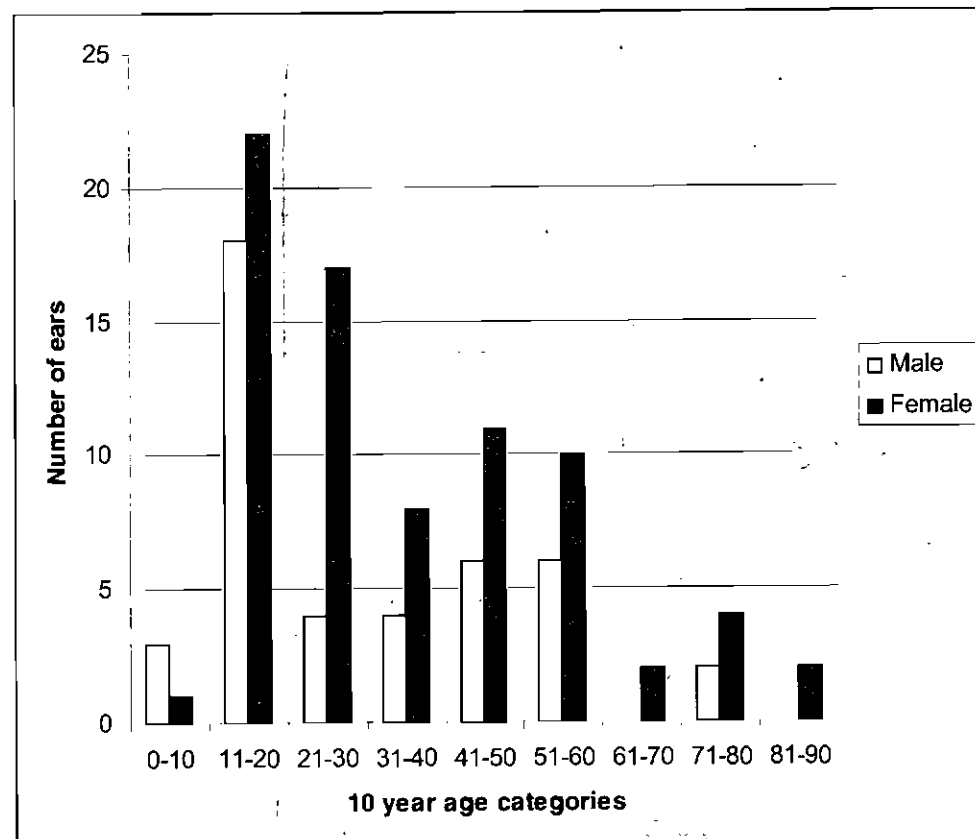


Figure 1: Age and gender representation of subjects in 10-year categories.

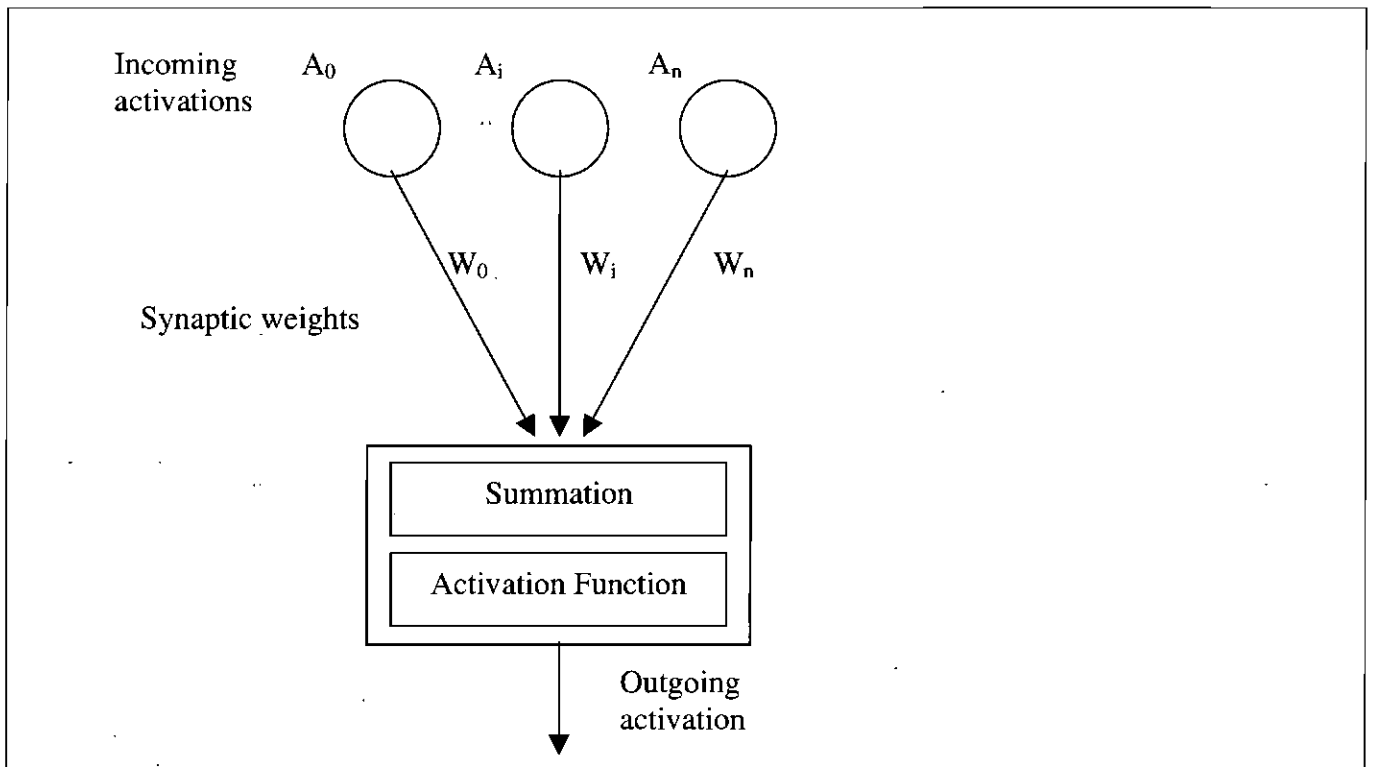


Figure 2: An Artificial neuron

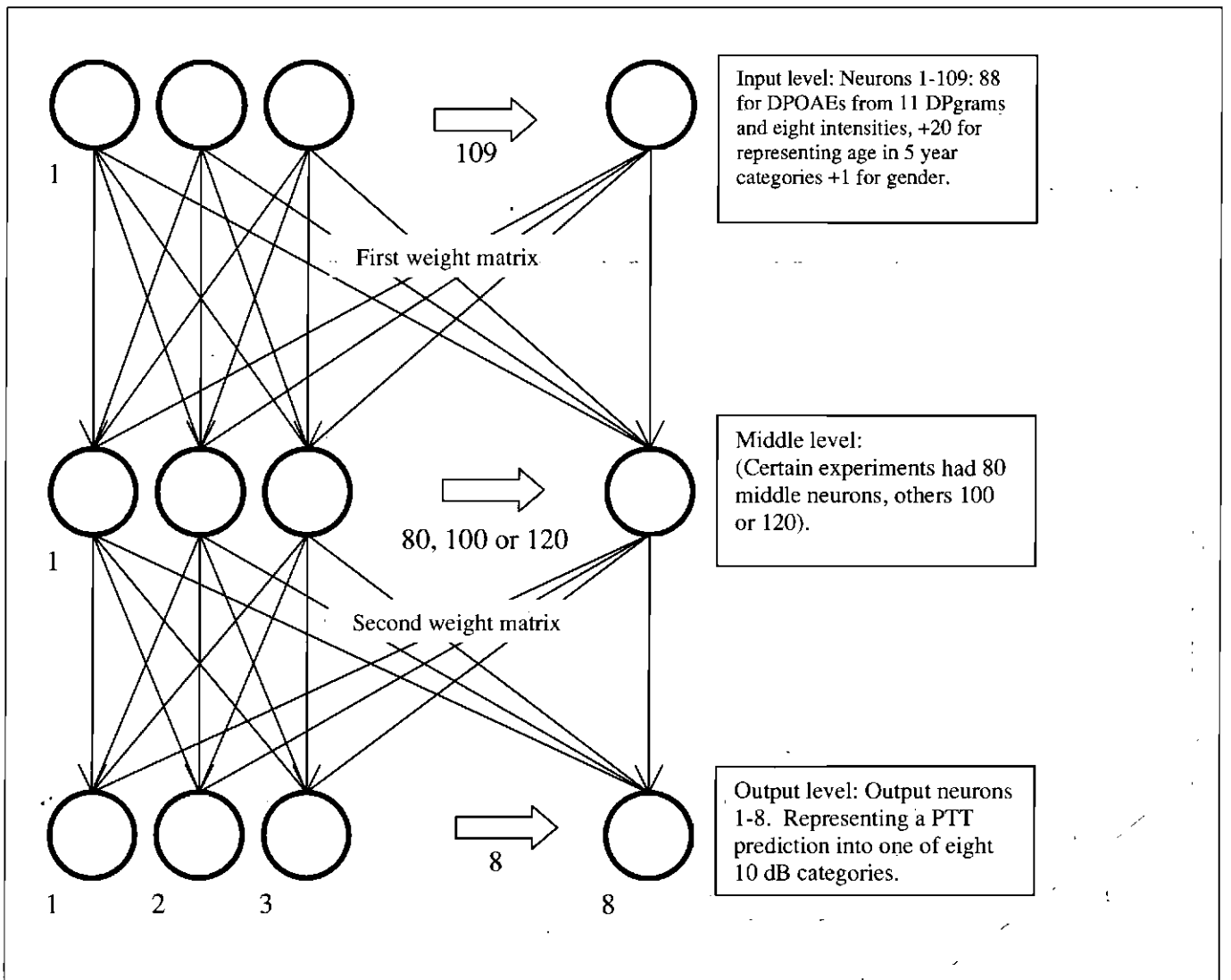


Figure 3: Schematic representation of a three-layered feed-forward artificial neural network

Reproduced by Sabinet Gateway under licence granted by the Publisher (dated 2012)

DATA PREPARATION AND PROCESSING

For this experiment, a commercially available feed-forward ANN with a back propagation training algorithm and one hidden layer was used, from Rao and Rao, (1995).

An artificial neural network is a computer program that consists of computational units, referred to as neurons. Neurons receive inputs analogous to the impulses that the dendrites of biological neurons receive, each input with its own mathematical value, or weight, indicating the importance of the input. The neuron calculates a total for all inputs, compares it to a threshold value and produces an output, just like a biological neuron sends an output through its axon. A schematic representation of an artificial neuron can be seen in Figure 2. Several neurons are combined to form an artificial neural network. A schematic representation of the three-layered type feed-forward network used in this study can be seen in Figure 3.

ANNs have two phases of operation, a training phase and a predicting phase. During training, the ANN assigns a weight (mathematical value) to every input it receives, and this weight affects the importance or impact of that input (Blum, 1992; Nelson & Illingworth, 1991). It is through repeated adjustments to the weights that the network learns. At first, weights are assigned at random. The ANN computes the output, compares it with the desired answer, and adjusts the weights repeatedly until the desired answers can be predicted for all the cases in the training set (Medsker, Turban, & Trippi, 1993). When prediction error for the training set is zero or acceptably low, the weights are frozen and the network is then presented with unfamiliar data to make a prediction based on the learned correlation (Rao & Rao, 1995). The topology of the network, such as the number of neurons in the hidden layer, error tolerance and learning rates were determined by experimentation but was not the result of an exhaustive search, so the possibility exists that different topologies might yield better results. In this study, there was experimentation with the number of hidden layer neurons as 80, 100 and 120 for all experiments except where amplitude was represented with the dummy variable technique as a categorical value (referred to as the ALT AMP technique). ALT AMP created much more inset neurons and the middle level neurons were doubled to compensate for additional neural network complexity. Acceptable error during training was 0.001, 0.002 and 0.003 (within 0.01%, 0.02% and 0.03% accurate). The Beta value or learning rate parameter was 0.5.

To prepare the DPOAE data for ANN training, all DPOAE responses from the eight DPgrams were rewritten in a binary fashion by using the dummy variable technique (Licht, 1998). Data was transformed into dichotomous variables indicating the presence or absence of a specific category. Absent DPOAE responses were depicted with a zero and present responses with a one. The pattern of all present and absent responses of the eight DPgrams served as input stimuli for the ANN. The ANN was presented with 88 input stimuli depicting DPOAE information of each ear (the eight DPgrams had 11 frequencies each, resulting in 88 possible DPOAE responses). The age variable was depicted in the same binary fashion as input into the ANN by indicating the 5

or 10-year category with ones and zeros: For example in the case of the 10-year category experiment, a 12-year old subject was depicted with a one in the second 10-year category and the other categories with zeros (01000000), an 82-year old subject as (00000001). Gender was depicted with a one or a zero. The amplitude of the DPOAE was either depicted as a fraction of the maximum DPOAE amplitude (40dB) measured in the study (referred to as AMP 40), a fraction of 100 (referred to as AMP 100) or by using the dummy variable technique by depicting it as a categorical value in one of four 10dB categories (referred to as ALT AMP).

For ANN training, DPOAE responses and PTTs of 118 ears were used to learn the correlation between DPOAEs and PTTs. Both ears of a subject were left out for each training phase. (The procedure was repeated for every subject, there were therefore 120 training phases.) In the prediction phase, the ANN was presented with only one of the DPOAEs of the two remaining "unfamiliar" ears. The ANN then predicted that ear's PTT based on the learned correlation of the training phase. This process was also repeated 120 times, to predict each ear once. The output of the ANN was a prediction of pure tone sensitivity at a specific frequency. This prediction however, was not in decibel form, but a categorization of hearing ability into eight 10dB categories.

Data analysis consisted of analyzing the actual and predicted values of all 120 ears and to determine how many were predicted accurately within the 10dB category, how many were predicted in an adjacent 10dB category and how many were predicted into a category more than 20dB out.

RESULTS

Tables 2 to 5 summarize results for the best neural network experiments predicting 500, 1000, 2000 and 4000 Hz. The tables present results in two categories: Correct predictions are classified as within the same 10dB category, one category out represents predictions made into an adjacent 10dB category. Remaining results were predicted more than one 10dB category out. Results are given for each 10dB category as well as overall ANN performance to predict normal hearing. The abbreviations for manipulated variables to outline the design of the specific experiment are explained in the key following each table.

PREDICTION OF NORMAL HEARING

The prediction of 500 Hz with DPOAEs has been problematic for many researchers due to the rising noise floor below 1000 Hz (Durrant, 1992; Gorga, et al., 1993; Stover et al., 1996). It seems that even normal and near normal ears exhibit no or small DPOAE amplitudes at 500 Hz (Probst & Hauser, 1990). In this study, normal hearing (<25 dB HL) at 500 Hz could be predicted with 94% accuracy. Normal hearing at 1000 Hz was predicted with 88% accuracy, 2000 Hz with 88% accuracy and 4000 Hz with 93% accuracy. The mean false positive rate for this study was 4% and the mean false negative rate was 16%. The false negative rate indicating the sensitivity of the procedure was still unacceptably high for diagnostic purposes (Brass & Kemp, 1994) even

Table 2: Prediction of 500 Hz with AI = 5, NoLF, Mid = 240, Err = 0.002, ALT AMP **

Catego-ries	1 (0 + 5dB)	2 (10 +15dB)	3 (20+25dB)	4 (30+35dB)	5 (40+45dB)	6 (50+55dB)	7 (60+65dB)	8 (70+75dB)
% Correct	80.9	9.7	0.0	16.7	22.2	0.0	0.0	*
% 10dB out	14.3	80.6	62.5	8.3	0.0	14.3	0.0	*
% False positive responses (0 – 25dBHL predicted as > 25 dBHL per category)				% False negative responses (> 25 dBHL predicted as 0 – 25dBHL per category)				
	0.0	2.0	1.0	6.0	5.0	4.0	0.0	*
ears in category	42	31	16	12	9	7	3	*
Overall prediction accuracy↓	**Key							
0-15dB predicted as 0-15dB	75%	AI = 5 → Age increment presented to ANN in 5-year categories						
0-15dB predicted as 0-25dB	95%	NoLF → No low frequency DPOAEs (<1000 Hz) included in ANN						
0-25dB predicted as 0-25dB	94%	ALT AMP → Amplitude of DPOAE presented as categorical value with dummy variable technique						
		Mid = 240 → Number of middle level neurons in ANN topology						
		Err = 0.002 → Error tolerance of ANN within 0.02% accurate						
* There were no ears in category eight, largest hearing loss measured for 500 Hz was 65dB HL.								

Table 3: Prediction of 1000 Hz with AI = 5, LF pres, Mid = 200, Err = 0.003, ALT AMP **

Catego-ries	1 (0 + 5dB)	2 (10 +15dB)	3 (20+25dB)	4 (30+35dB)	5 (40+45dB)	6 (50+55dB)	7 (60+65dB)	8 (70+75dB)
% Correct	80.0	24.1	33.3	0.0	20.0	0.0	18.2	*
% 10dB out	15.0	55.2	25.0	20.0	6.7	12.5	9.1	*
% False positive responses (0 – 25dBHL predicted as > 25 dBHL per category)				% False negative responses (> 25 dBHL predicted as 0 – 25dBHL per category)				
	0.0	5.0	1.0	2.0	10.0	5.0	5.0	*
ears in category	40	29	12	5	15	8	11	*
Overall prediction accuracy↓	**Key							
0-15dB predicted as 0-15dB	86%	AI = 5 → Age increment presented to ANN in 5-year categories						
0-15dB predicted as 0-25dB	89%	LF pres → Low frequency DPOAEs (<1000 Hz) included in ANN						
0-25dB predicted as 0-25dB	88%	ALT AMP → Amplitude of DPOAE presented as categorical value with dummy variable technique						
		Mid = 200 → Number of middle level neurons in ANN topology						
		Err = 0.003 → Error tolerance of ANN within 0.03% accurate						
* There were no ears in category eight, largest hearing loss measured for 1000 Hz was 65dB HL.								

though it was lower than reported elsewhere (Gorga et al., 1993; Kimberley et al., 1994b; Stover et al., 1996).

Reasons for the improved prediction of normal hearing at low frequencies involve a number of possibilities: First, it might be possible that the ANN found significant information in the fine structure of the higher frequency DPOAEs hidden in the pattern of all present and absent responses of the eight DPgrams to enable an accurate prediction at 500 Hz. Another

possibility is the fact that PTTs are heavily interrelated and with enough information at the high frequencies, it is possible for the ANN to predict the low frequency as one of a limited number of “audiogram pattern” or “DPOAE pattern” possibilities. It should be mentioned however, that this neural network prediction at 500 Hz is not mere coincidence. Neural networks cannot do magic, a clear-cut correlation is needed to enable an accurate prediction. Two unrelated data sets will produce predictions that

approximate random noise. The ANN was therefore able to extract enough information from the DPOAE spectrum to correctly classify normal hearing at 500 Hz with 94% accuracy.

normal or impaired as low as 500 Hz was surprisingly accurate, predictions of categories representing hearing loss were disappointingly poor. Table 6 summarizes overall prediction accuracy across all eight 10dB categories for all frequencies.

PREDICTION OF HEARING LOSS

Even though classification of hearing ability as

Although prediction of PTTs within a 10 dB category representing hearing loss was correct only up to 32 % to 40% of the time, prediction into an adjacent

Table 4: Prediction of 2000 Hz with AI = 5, LF pres, Mid = 80, Err = 0.002, No AMP **

Categories	1 (0 + 5dB)	2 (10 +15dB)	3 (20+25dB)	4 (30+35dB)	5 (40+45dB)	6 (50+55dB)	7 (60+65dB)	8 (70+75dB)
% Correct	78.1	41.4	0.0	0.0	0.0	6.3	7.7	*
% 10dB out	15.6	55.2	60.0	22.2	45.5	56.3	30.8	*
% False positive responses (0 – 25dBHL predicted as > 25 dBHL per category)				% False negative responses (> 25 dBHL predicted as 0 – 25dBHL per category)				
	1.0	0.0	4.0	4.0	4.0	4.0	3.0	*
ears in category	32	29	10	9	11	16	13	*
Overall prediction accuracy ↓	**Key							
0-15dB predicted as 0-15dB	90%	AI = 5 → Age increment presented to ANN in 5-year categories LF pres → Low frequency DPOAEs (<1000 Hz) included in ANN No AMP → Amplitude of DPOAE omitted in ANN training Mid = 80 → Number of middle level neurons in ANN topology Err = 0.002 → Error tolerance of ANN within 0.02% accurate						
0-15dB predicted as 0-25dB	95%							
0-25dB predicted as 0-25dB	88%							
* There were no ears in category eight, largest hearing loss measured for 2000 Hz was 65dB HL.								

Table 5: Prediction of 4000 Hz with AI = 5, LF pres, Mid = 80, Err = 0.002, No AMP**

Categories	1 (0 + 5dB)	2 (10 +15dB)	3 (20+25dB)	4 (30+35dB)	5 (40+45dB)	6 (50+55dB)	7 (60+65dB)	8 (70+75dB)
% Correct	84.8	27.8	0.0	12.5	0.0	45.0	22.2	11.1
% 10dB out	15.2	61.1	57.1	0.0	28.6	30.0	33.3	44.4
% False positive responses (0 – 25dBHL predicted as > 25 dBHL per category)				% False negative responses (> 25 dBHL predicted as 0 – 25dBHL per category)				
	0.0	1.0	1.0	3.0	2.0	2.0	5.0	1.0
ears in category	33	18	7	8	7	20	18	9
Overall prediction accuracy ↓	**Key							
0-15dB predicted as 0-15dB	92%	AI = 5 → Age increment presented to ANN in 5-year categories LF pres → Low frequency DPOAEs (<1000 Hz) included in ANN No AMP → Amplitude of DPOAE not included in ANN training Mid = 80 → Number of middle level neurons in ANN topology Err = 0.002 → Error tolerance of ANN within 0.02% accurate						
0-15dB predicted as 0-25dB	96%							
0-25dB predicted as 0-25dB	93%							

Table 6: Overall prediction accuracy across all eight 10 dB categories for 500, 1000, 2000 and 4000 Hz.

	Correct (Within 10dB)		Predicted into an adjacent 10dB Category		More than one 10dB category out	
	ears	%	ears	%	ears	%
500 Hz	41/120	34.2	44/120	36.6	35/120	29.2
1000 Hz	48/120	40.0	29/120	24.2	43/120	35.8
2000 Hz	39/120	32.0	47/120	39.0	34/120	28.0
4000 Hz	48/120	40.0	38/120	31.0	34/120	28.0

10dB category added up to an additional 24% to 39%, and therefore provided an indication of severity of hearing loss even though it was a less specific classification of hearing.

There are several possibilities to explain the poor prediction of categories representing hearing loss. The first reason that can be suggested for the poor

prediction of hearing impaired categories is that hearing impaired subjects might demonstrate less clear DPOAE responses that might influence the correlation between DPOAEs and PTTs. This possibility was seen in the DPOAE evaluation where it took the DPOAE equipment longer to test a hearing impaired subject in order to get enough frames of data to regard a test as "accepted."

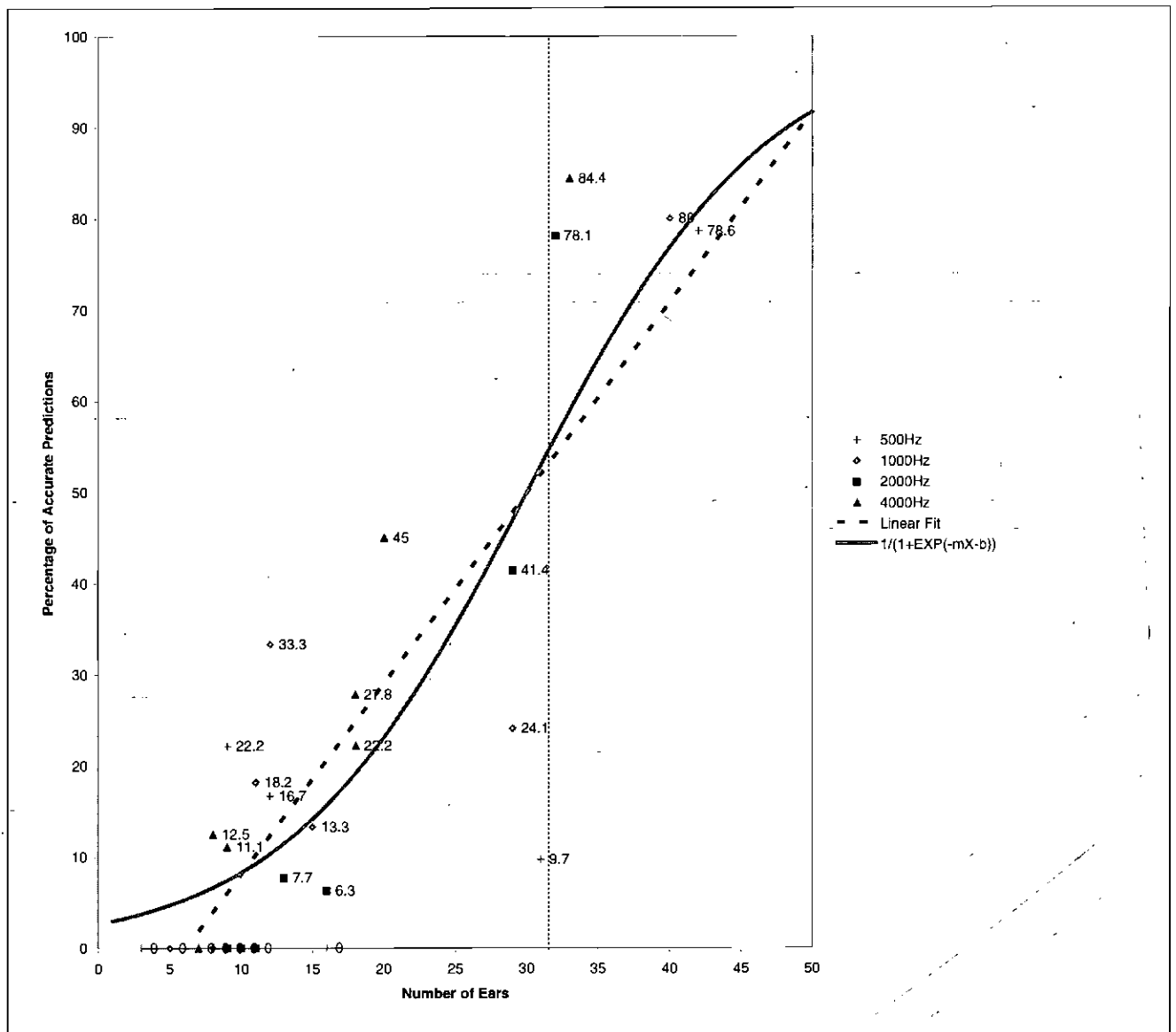


Figure 4: Prediction accuracy and ear count correlation.

Reproduced by Sabinet Gateway under licence granted by the Publisher (dated 2012)

(The criterion for a test to be accepted for the GSI-60 DPOAE system for a screening test is that the DPOAE amplitude had to be 10dB above the noise floor or the cumulative noise level had to be -18dB SPL. The maximum number of frames tested in a screening procedure is 400, and if no clear response is measured in that time, the test is scored "timed out" which means that no response was obtained.) It is possible that more responses could be obtained from hearing-impaired subjects if the criterion for test acceptance is lowered to 5 dB for example. The lowered criterion for the acceptance of a test could possibly enhance the number of useable responses from hearing-impaired subjects and might therefore enhance prediction accuracy of categories spanning impaired hearing. This aspect should be further investigated.

Another reason for poor prediction accuracy of hearing impaired categories might be that the optimal procedure for data analysis has not yet been identified. It can be hypothesized that another type of neural network with different topologies, learning rules and error tolerances would be able to make more accurate predictions. Another possibility is a complete new form of data processing, such as "genetic programming", inspired by Darwinian invention and problem solving that "progressively breeds a population of computer programs over a series of generations" (Koza, Bennett, Andre, & Keane, 1999, p.3) to find an optimal solution to a problem. These possibilities have yet to be investigated. One definite aspect however, that seems to have influenced the prediction accuracy of hearing ability in categories spanning hearing impairment more than neural network capabilities and the underlying correlation between the two data sets, was the number of ears in every category that the neural network had to train on.

Neural networks need enough examples in every category to form representations of how a specific ear's DPOAE type relate to its PTT type in order to make accurate predictions. Even though this study attempted to categorize all audiograms into three groups to ensure that hearing impairment was as well represented as normal hearing, the nature of sensorineural hearing impairment tends to affect certain frequencies more than others, and low frequencies are often normal. In the case of 500 Hz, this leads to the uneven distribution of many ears in categories representing normal hearing and few ears in hearing loss categories. At 4000 Hz however, category six (50 + 55dB) had more ears (a total of 20) and was predicted accurately 45% of the time and within 10dB 30% of the time. Hearing loss at 4000Hz in this category was predicted incorrectly only 25% of the time with a false negative rate of only 2%. The same category at 500 Hz had only 7 ears to train on and prediction accuracy was never correct and within 10dB only 14% of the time. If the network had more ears in every category to train on, prediction accuracy might have been considerably better.

To illustrate this concept, Figure 4 plots the number of ears in every category against prediction accuracy. It clearly demonstrates that the number of ears in every category had an enormous effect on prediction accuracy.

The aim of this study is not to describe the relationship between the number of ears per category and prediction accuracy thereof. However, Figure 4 shows a

number of possible relationships merely to illustrate the notion of "more-is-better". The linear fit (in dotted lines) shows that higher numbers of ears lie significantly above expectation. It is worth noting that the relationship cannot be linear, since any line with a slope larger than zero will have to cross the 100% accuracy limit at some point, which is of course not possible.

The figure indicates an alternative fit (the solid line) of the form $1/(1 + e^{-mx-b})$, that seems better and more intuitive since it starts out low, just like the experimental data, but also asymptotically approaches 100% for large numbers of ears. It has also not been established if this function has any correlation with the sigmoid function (Blum, 1992:39) $(1 / (1 + e^{-x}))$ that was used to normalize output of the separate ANN layers but it seems to be a more likely option than a linear fit.

For this data set there is a fairly clear threshold at 32 ears per category where prediction accuracy suddenly surges into the 75% and higher region. Should any of the example relationships in Figure 4 hold, it is expected that a 95% or higher accurate predictions could potentially be made if an ANN receives 80 or more ears per category.

CONCLUSION

ANNs were able to subtract significant information from the DPOAE spectrum to form an internal representation of the correlation between DPOAEs and PTTs and used that information successfully to predict normal hearing with high levels of accuracy as low as 500 Hz. The unsatisfactory prediction of categories representing hearing loss is most likely due to the shortage of data in certain categories and not poor correlation between DPOAEs and PTTs of hearing-impaired ears or incapability of the neural network to deal with this data set. With more data, it seems possible to predict PTTs within 10 dB from 500 Hz to 4000 Hz, for hearing losses up to 65dB HL.

REFERENCES

- Avan, P., & Bonfils, P. (1993). Frequency specificity of human distortion product Otoacoustic Emissions. *Audiology*, 32, 12-26.
- Blum, A. (1992). *Neural networks in C++: An object-oriented framework for building connectionist systems*, NY: John Wiley & Sons.
- Bonfils, P., Avan, P., Londero, A., Trotoux, J., & Narcy, P. (1991). Objective low-frequency audiometry by distortion-product acoustic emissions. *Archives of Otorhinolaryngology-Head and Neck Surgery*, 117, 1167-1171.
- Brass, D., & Kemp, D.T. (1994). Quantitative assessment of methods for the detection of otoacoustic emissions. *Ear and Hearing*, 15, 378-389.
- De Waal, R. (1998). The use of artificial neural networks to predict pure tone thresholds in normal and hearing-impaired ears with distortion product otoacoustic emissions. Unpublished master's thesis, University of Pretoria, South Africa.
- De Waal, R. (2000). Objective prediction of pure tone thresholds in normal and hearing-impaired ears with distortion product otoacoustic emissions

- and artificial neural networks. Unpublished doctoral dissertation, University of Pretoria, South Africa.
- Dorn, P.A., Piskorski, P., Gorga, M.P., Neely, S.T., & Keefe, D.H.** (1999). Predicting audiometric status from distortion product otoacoustic emissions using multivariate analyses. *Ear and Hearing, 20*, 149-163.
- Durrant, J.D.** (1992). Distortion-product OAE analysis: Is it ready for broad clinical use? *The Hearing Journal, 45*(11), 42-45.
- Gaskill, S.A., & Brown A.M.** (1990). The behavior of the acoustic distortion product, 2f1-f2, from the human ear and its relation to auditory sensitivity. *Journal of the Acoustical Society of America, 88*(2), 821-839.
- Gorga, M.P., Neely, S.T., Bergman, B., & Beauchaine, K.L.** (1993). Otoacoustic emissions from normal-hearing and hearing-impaired subjects: Distortion product responses. *Journal of the Acoustical Society of America, 93*(4), 2050-2060.
- Grason-Stadler Inc. User manual 1760-0100, Rev 8 for the GSI-60 DPOAE.** (1997). Grason-Stadler, Inc, NH 3055-3056.
- Harris, F.P., Lonsbury-Martin, B.L., Stagner, B.B., Coats, A.C., & Martin, G.K.** (1989). Acoustic distortion products in humans: Systematic changes in amplitude as a function of f2/f1 ratio. *Journal of the Acoustical Society of America, 85*(1), 220-229.
- Harris, F.P., & Probst, R.** (1991). Reporting click-evoked and distortion-product otoacoustic emission results with respect to the pure-tone audiogram. *Ear and Hearing, 12*(6), 399-405.
- He, N., & Schmiedt, R.** (1996). Effects of aging on the fine structure of the 2f1-f2 acoustic distortion product. *Journal of the Acoustical Society of America, 99*(2), 1002-1015.
- Hurley, R.M., & Musiek, F.E.** (1994). Effectiveness of transient-evoked otoacoustic emissions (TEOAEs) in predicting hearing level. *Journal of the American Academy of Audiology, 5*(3), 195-203.
- Kalluri, R., & Shera, A.S.** (2001). Distortion-product source mixing: A test of the two-mechanism model for DPOAE generation. *Journal of the Acoustical Society of America, 109*(2), 622-637.
- Karzon, R.K., Garcia, P., Peterein, J.L., & Gates, G.A.** (1994). Distortion product otoacoustic emissions in the elderly. *The American Journal of Otology, 15*(5), 596-605.
- Kemp, D.T.** (1978). Stimulated acoustic emissions from within the human auditory system. *Journal of the Acoustical Society of America, 64*, 1386-1391.
- Kimberley, B.P., Hernadi, I., Lee, A.M., & Brown, D.K.** (1994b). Predicting pure tone thresholds in normal and hearing-impaired ears with distortion product emission and age. *Ear and Hearing, 15*(3), 199-209.
- Kimberley, B.P., Kimberley, B.M., & Roth, L.** (1994a). A neural network approach to the prediction of pure tone thresholds with distortion product emissions. *ENT Journal, 73*(11), 812-823.
- Kimberley, B.P., & Nelson, D.A.** (1989). Distortion product emissions and sensorineural hearing loss. *The Journal of Otolaryngology, 18*(7), 365-369.
- Knight, R.D., & Kemp, D.T.** (1999a). Indications of different distortion product otoacoustic emission mechanisms from a detailed f1, f2 area study. *Journal of the Acoustical Society of America, 107*(1), 457-473.
- Koza, J.R., Bennett, F.H., Andre, D., & Keane, M.A.** (1999). *Genetic Programming III*. San Francisco: Morgan Kaufmann Publishers.
- Kummer, P., Janssen, T., & Arnold, W.** (1998). The level and growth behavior of the 2f1 - f2 distortion product otoacoustic emission and its relationship to auditory sensitivity in normal hearing and cochlear hearing loss. *Journal of the Acoustical Society of America, 103*(6), 3431-3444.
- Lee, A., Kimberley, B., & Brown, D.** (1993). Developing distortion product emission measurements for clinical applications. *Journal of Speech Language Pathology and Audiology, 17*(1), 15-25.
- Licht, M.H.** (1998). Multiple regression and correlation. In G.L. Grimm and P.R. Yarnold, (Eds.), *Reading and understanding multivariate statistics* (pp. 19-64). Washington, DC: American Psychological Association.
- Lonsbury-Martin, B.L., Cutler, W.M., & Martin, G.K.** (1991). Evidence for the influence of aging on distortion product emissions in humans. *Journal of the Acoustical Society of America, 89*, 1749-1759.
- Lonsbury-Martin, B.L., & Martin, G.K.** (1990). The clinical utility of distortion-product otoacoustic emissions. *Ear and Hearing, 11*(2), 144-154.
- Martin, G.K., Ohlms, L.A., Franklin, P.H., Harris, F.P., & Lonsbury-Martin, B.L.** (1990). Distortion product emissions in humans III. Influence of sensorineural hearing loss. *Annals of Otology, Rhinology and Laryngology, 99*, 30-42.
- Mauermann, M., Uppenkamp, S., Van Hengel, P.J.W., & Kollmeier, B.** (1999a). Evidence for the distortion product frequency place as a source of distortion product otoacoustic emission (DPOAE) fine structure in humans. I. Fine structure and higher-order DPOAE as a function of the frequency ratio f2/f1. *Journal of the Acoustical Society of America, 106*(6), 3473-3483.
- Mauermann, M., Uppenkamp, S., Van Hengel, P.J.W., & Kollmeier, B.** (1999b). Evidence for the distortion product frequency place as a source of distortion product otoacoustic emission (DPOAE) fine structure in humans. II. Fine structure for different shapes of cochlear hearing loss. *Journal of the Acoustical Society of America, 106*(6), 3484-3491.
- Medsker, L., Turban, E., & Trippi, R.** (1993). Neural network fundamentals for financial analysts. In R.R. Trippi and E. Turban (Eds.), *Neural networks in finance and investing: Using artificial neural networks to improve real-world*

- performance (pp.4-49). Chicago: Probus Publishing Company.
- Mills, D.M.** (1997). Interpretation of distortion product otoacoustic emission measurements I. Two stimulus tones. *Journal of the Acoustical Society of America*, 102(1), 413-429.
- Moulin, A., Bera, C., & Collet, L.** (1994). Distortion product otoacoustic emissions and sensorineural hearing loss. *Audiology*, 33, 305-326.
- Musiek, F.E., Berenstein, S.P., Hall III, J.W., & Schwaber, M.K.** (1994). Auditory brainstem response: Neurodiagnostic and intraoperative applications. In J. Katz (Ed.), *Handbook of clinical audiology* (4th ed.), (pp. 351-374). London: Williams & Willkins.
- Nakajima, H.H., Mountain, D.C., & Hubbard, A.E.** (1998). Nonlinear characteristics of electrically evoked otoacoustic emissions. *Hearing Research*, 122, 109-118.
- Nelson, M.M., & Illingworth, W.T.** (1991). *A practical guide to neural nets*. New York, NY: Addison-Wesley Publishing Company, Inc.
- Nieschalk, M., Hustert, B., & Stoll, W.** (1998). Distortion product otoacoustic emissions in middle-aged subjects with normal versus potentially presbycusis high-frequency hearing loss. *Audiology*, 37, 83-99.
- Northern, J.L.** (1991). *Hearing in children*. Baltimore, MD: Williams & Willkins.
- Ohlms, L.A., Lonsbury-Martin, B.L., & Martin, G.K.** (1990). Acoustic-distortion products: Separation of sensory from neural dysfunction in sensorineural hearing loss in human beings and rabbits. *Otolaryngology-Head and Neck Surgery*, 104(2), 159-174.
- Probst, R., & Hauser, R.** (1990). Distortion product otoacoustic emissions in normal and hearing-impaired ears. *American Journal of Otolaryngology*, 11, 236-243.
- Rao, V., & Rao, H.** (1995). *C++ neural networks and fuzzy logic* (2nd ed.). NY: MIS Press.
- Robinette, M.S.** (1992). Otoacoustic emissions in cochlear vs. retrocochlear auditory dysfunction. *The Hearing Journal*, 45(11), 32-34.
- Smurzynski, J., Leonard, G., Kim, D.O., Lafreniere, D.C., Marjorie, D., & Jung, M.D.** (1990). Distortion product otoacoustic emissions in normal and impaired adult ears. *Archives of Otolaryngology and Head and Neck Surgery*, 116, 1309-1316.
- Stover, L., Gorga, M.P., & Neely, S.T.** (1996). Toward optimizing the clinical utility of distortion product otoacoustic emission measurements. *Journal of the Acoustical Society of America*, 100(2), 956-967.
- Talmadge, C.L., Long, G.R., Tubis, A., & Dhar, S.** (1998). Experimental confirmation of the two-source interference model for the fine structure of distortion product otoacoustic emissions. *Journal of the Acoustical Society of America*, 105(1), 275-292.
- Tanaka, Y., O-Uchi, T., Arai, Y., & Suzuki, J.** (1987). Otoacoustic emission as an indicator in evaluating inner ear impairments. *Acta Otolaryngologica, (Stockholm)* 103, 644-648.
- Weber, B.A.** (1994). Auditory brainstem response: Threshold estimation and auditory screening. In J. Katz (Ed.), *Handbook of Clinical Audiology*. (4th ed.), (pp. 375-386). London: Williams & Willkins.
- Yantis, P.A.** (1994). Puretone air-conduction threshold testing. In J. Katz (Ed.), *Handbook of Clinical Audiology*. (4th ed.), (pp. 97-108). London: Williams & Willkins.
- Zhao, F., & Stephens, D.** (1998). Analysis of notches in audioscan and DPOAEs in subjects with normal hearing. *Audiology*, 37, 335-343.